学校编码：10384　　分类号_____密级_____
学号：20520091151329　　UDC___

厦门大学

硕 士 学 位 论 文

# 蛋白质多维分析的新型多功能探针的设计合成和标记策略研究

# Design and Synthesis and Labeling Strategy of Novel Probes for Multidimensional Analysis of Proteins

张 珠 宝

指导教师姓名：　王 秋 泉　教授
专 业 名 称：　分 析 化 学
论文提交日期：　2 0 1 2 年　　月
论文答辩时间：　2 0 1 2 年　　月
学位授予日期：　2 0 1 2 年　　月

答辩委员会主席：_____
评　　阅　　人：_____

2012　年　月

# 厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下，独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（　　）课题（组）的研究成果，获得（　　）课题（组）经费或实验室的资助，在（　　）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）：

年　　月　　日

# 厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（　　）1.经厦门大学保密委员会审查核定的保密学位论文，于　　年　　月　　日解密，解密后适用上述授权。

（　　）2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年　　月　　日

# 目 录

# CONTENTS

# 摘 要

继基因组学之后，蛋白质组学成为了当今生命科学、化学和医学科学研究中共同关注的另一大课题和研究热点。随着蛋白质组学研究的深入，对同一生物体系中蛋白质表达种类的识别和鉴定、表达位置及分布的定位分析，以及更重要的其表达水平动态变化的准确定量，以便获得全面而准确的信息，已经成为了蛋白质组学的重要研究目标和急需解决的难题，对生命奥秘的揭示、临床诊断以及药物筛选等领域的发展具有重要的意义。在极具挑战的蛋白质组学研究的进程中，发展高灵敏、高选择性的新分析技术和方法是解决复杂生物体系中蛋白质分析所面临的新问题的有效途径，各国的分析科学家和生物学家等也不断地朝这个方向开展深入的研究。在近年来所发展的多种分析技术和方法中，基于标记策略的荧光和质谱技术成为了蛋白质组学研究中的两大主流技术。荧光技术具有灵敏度高（尤其是 LIF 是最灵敏的检测方法）、时空分辨能力强；荧光成像可实现可视化，能够对蛋白质进行直观的实时和原位观察和分析。而质谱技术，分子质谱具有强大的结构分析能力，能够鉴定蛋白质的种类并可提供相对定量的信息；作为元素最理想的分析技术—元素质谱（ICP-MS）具有其他技术无法比拟的准确和绝对定量的能力，通过蛋白质含有的内源元素或者外源元素的标记，结合专一的识别方法或者高效的分离技术，能够实现蛋白质的准确绝对定量。这些技术各有优势和不足，相对于单一技术的应用，若将荧光和质谱技术有机结合起来将为蛋白质组学的研究提供一个更加强大和相对完善的分析平台，将为解决同一样品中蛋白质的识别、可视化和量化等难题提供可靠的新途径和有力的新工具，推动蛋白质组学研究的发展。

针对蛋白质研究的选择性识别、可视化定位和准确定量等同时分析的要求，我们发展了新型的多功能探针和标记策略，开展了蛋白质的荧光和元素质谱联合检测分析方法研究。本硕士论文全文主要包括以下四个部分：

第一章　对蛋白质的背景知识进行了概述，着重介绍和综述了蛋白质的各种标记策略和相应的分析检测技术和方法的研究进展，在此基础上提出了本论文的选题依据、研究内容和意义。

第二章　发展了基于稀土元素标签MMA-DOTA-Eu和荧光标签FITC依次标

记同一目标分子于蛋白分子的活性巯基和氨基的双标记策略，并建立了基于这种双标记策略的多肽/蛋白质柱后同位素稀释-元素质谱 ICP-MS 绝对定量和激光诱导荧光 LIF 检测的二维分析平台；并成功应用于三种模型肽 Vas、Som 和 GGYGGC 的检测中，实现了三种多肽的高灵敏识别、分离和定量分析。通过比较结果可知，HPLC-ICP-MS 对三种多肽的检出限分别为 0.130，0.130 和 0.259 nmol $L^{-1}$，比 CE-LIF 低出两个数量级，显示出 HPLC-ICP-MS 在多肽/蛋白质的检测中的优势。该方法初步实现了对同一目标分子的荧光和元素质谱同时检测的目标。

第三章　设计合成了集荧光基团、元素基团和靶向基团于一体的三功能探针 cRGD-AMF-DOTA-Eu，并结合激光共聚焦显微镜（CLSM）和元素质谱(ICP-MS)等技术，建立了基于靶向多肽介导的多功能探针的癌症标志物和癌细胞的荧光成像和 ICP-MS 检测的新方法。我们将其应用于癌症的重要标志物$\alpha_v\beta_{3/5}$整联蛋白和相关的癌细胞的研究，同时实现了它们的高特异性标记、高灵敏识别、高分辨荧光成像定位与准确定量，获得可视化和量化等更加全面和准确的信息，为癌症的早期诊断开辟了一条新途径。

第四章　总结了本论文的研究工作，并对这一领域将来进一步的发展和研究趋势进行了展望。

关键词：蛋白质；双标记；三功能标签；电感耦合等离子体质谱；同位素稀释；蛋白质绝对定量；毛细管电泳–激光诱导荧光检测；荧光成像；癌细胞；癌症早期诊断

# Abstract

After genomics, proteomics has become another major issue and thus a research focus in life science and chemical biology and medical science. As the deepening of the proteomics research, besides to recognize the species and to visualize their expression location and distribution, it is more important to accurately quantify the expression level during the dynamic change of proteins presented in the same given biological sample, obtaining the more comprehensive and accurate information, which is of great significance to understand the mysteries of life, and to perform a reliable clinical diagnosis as well as subsequent drug screening. In order to effectively solve these issues emerged in proteomics studies, new analytical techniques and methods of high sensitivity and selectivity are urgently needed. Among various analytical techniques and methods developed in recent years, fluorescence and mass spectrometry (MS) based on labeling strategies have become the two main analytical techniques for proteomics research. The power of fluorescence technique lies in its high sensitivity (especially, LIF, which is one of the most sensitive spectroscopic methods), high spatial and temporal resolution, visual imaging ability, thus can realize the real-time and in-situ monitoring of proteins in vitro and in vivo. As for MS, molecular MS is a very effective tool for structural identification of proteins in the biological system and also for their relative quantification. Element MS (especially, ICP-MS), which is the best element measuring technique so far available, can achieve the absolute protein quantification via measuring the endogenous heteroatoms in the protein molecules and/or exogenous elements introduced onto the proteins by labeling strategies when coupled with a protein-specific recognition route or an effective separation technique. Taking advantage of the strengths of the above techniques and overcoming their individual limitations, the combination of MS and fluorescence spectrometry is probably a most powerful way for recognizing, visualizing and accurately quantifying peptides/proteins in biological samples, giving comprehensive information of targets, paving a reliable new way for proteomics research and promoting the rapid development of proteomics.

In this thesis, in order to selectively recognize, visualize and accurately quantify proteins in the same sample, we developed several novel labeling strategies based on the develpement of several mono-function probes or one multifunctional probe, which allow multi-technique analysis of proteins combining the use of fluorescence detection technique and element mass spectrometry. These will be described in the below four chapters:

In Chapter 1, background knowledge of proteins was briefly introduced first. Furthermore, recent development of protein labeling strategies and the corresponding detection techniques and analytical methods were emphatically described and reviewed. On this basis, the research ideas and contents and significance of this thesis were proposed.

In Chapter 2, a new dual-labeling strategy for labeling -SH and -$NH_2$ of the same target peptide molecule with a rare-earth element tag (MMA-DOTA-Eu) and a fluorescent tag FITC was developed. Based on this strategy, a two-dimensional analysis platform for peptide detection and absolute quantification was developed using species-unspecific isotope dilution strategy with ICP-MS and LIF techniques. This method successfully applied to the analysis of the three model peptides Vas, Som and GGYGGC, realizing their highly sensitive recognization, separation and quantification. As can be seen from the results, the LODs of the three model peptides obtained using HPLC-ICP-MS were two orders of magnitude lower than those using CE-LIF, reaching 0.130，0.130 and 0.259 nmol $L^{-1}$ for Vas, Som and GGYGGC, respectively, suggesting that HPLC-ICP-MS is a superior platform for quantification of peptides.

In Chapter 3, a new multidimensional analytical method using laser confocal scanning microscope (CLSM) and ICP-MS for the quantification of cancer markers and the corresponding cancer cells was established based on the design and synthesis of a trifunctional probe cRGD-AMF-DOTA-Eu, in which fluorophore group, element moiety and guiding group were conjugated together. It was applied to study the important cancer markers $\alpha_v\beta_{3/5}$ integrins and the corresponding cancer cells, achieving their highly specific labeling, high-resolution fluorescence imaging and

accurate quantification. More comprehensive and accurate information of the biomarker and corresponding tomer cells were obtained, paving a new way for the early diagnosis of cancer.

In Chapter 4, a conclusion of what I have achieved during my master studies and a perspective of future development in this research field were presented.

Keywords: Protein; Dual-labeling Strategy; Trifunctional Probe; ICP-MS; Isotope Dilution; Protein Absolute Quantification; CE-LIF; Fluorescence Imaging; Cancer Cell; Early Diagnosis of Cancer

# 第一章 前言

## 1.1 蛋白质的概述

近年来，随着人类基因组精确图谱的完成，生命科学已经进入后基因组时代。基因是遗传信息的携带者，在实际生命活动中基因调控着蛋白质的合成和表达。蛋白质是生物体内含量最高，功能最重要的生物大分子，它是生命的重要物质基础，与各种生命活动都是紧密地联系着，参与着生物体内的每一个细胞和所有的组成部分，扮演着构筑生命大厦不可缺少的重要角色，它们在化学组成、分子结构、生物功能等方面具有有别于其他生物分子的特点，对蛋白质的研究有利于帮助人们从分子水平探讨和了解生命活动的规律和重要的生理、病理现象的本质。随着研究的深入，人们了解到基因组的研究成果的获得能够为疾病提供依据，然而基因组测序只是迈出了认识复杂生命体的万里长征的第一步，与其表达产物蛋白质比较起来，它远没有蛋白质与疾病的相关性大，且很多问题无法从基因水平寻找到答案。因此，生命科学家将生命科学的研究目标转移到了蛋白质组学，蛋白质领域的研究已经成为了 21 世纪科学领域的焦点之一[1-3]。

蛋白质是由 20 种基本的氨基酸组成的一条或者多条肽链以特殊方式组合的生物大分子化合物，其复杂结构主要包括一级、二级、三级或四级结构[4]，而且这些结构在不同的化学环境和条件下还会发生相应的变化，从而达到改变其化学反应活性并调节生物化学反应平衡的目的[5]。研究表明，蛋白质是生命体功能的最终执行者，在分子水平和细胞水平参与着各种生理和病理过程，发挥着各自不同的功能。蛋白质的表达与否、表达多少等等受生命生长和外部环境等多种因素的影响，不同的细胞、组织、器官以及甚至在同一位置在正常和疾病状态下它们的表达水平和存在形态也各不相同。因而，蛋白质的含量、种类、翻译后的修饰以及在细胞、组织和体内位置的变化决定和反映了相应的生理功能的变化，揭示着生命过程的变化，其这些方面的异常预示着某种疾病的可能发生和病变发展阶段，如癌症、白血病等等。众所周知，攻克疾病特别是癌症的最有效的方法是及早发现发病的蛛丝马迹，实现早期诊断并及时治疗，而科学家发现每种疾病可激发出一套它所独有的蛋白质“生物标记”，所以这一目标的实现依赖于对这些特